

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

(54) JAPANESE LANGUAGE SYNTAX ANALYZING SYSTEM

(11) 4-184677 (A) (43) 1.7.1992 (19) JP

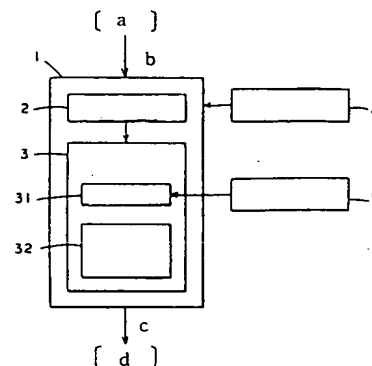
(21) Appl. No. 2-315229 (22) 20.11.1990

(71) CSK CORP (72) AKIRA TAKAGI(1)

(51) Int. Cl. G06F15/38

PURPOSE: To accurately decide the kind of a participial modification clause by attaching meaning constraint information on a word with indefinite meaning based on case relation between the word with indefinite meaning and a following predicate, and judging the case relation between the word with indefinite meaning and the predicate in the participial modification clause based on the meaning constraint information.

CONSTITUTION: The morpheme analysis part 2 of an analysis processor 1 divides an inputted Japanese language text into every word, and attaches grammar information and meaning information on each word by referring to a Japanese language dictionary 4. Thence, a syntax analysis part 3 judges the modification relation of divided words based on the meaning information, the grammar information, and part of speech matrix information, and analyzes structure. At this time, when the word modified by the participial modification clause is the one with indefinite meaning, a modification judging part 31 judges the case relation between the following predicate modified by the word with indefinite meaning, and attaches the meaning constraint information on the word with indefinite meaning. Thence, a participial modification clause decision part 32 judges the case relation between the word with indefinite meaning and the predicate in the participial modification clause. Thereby, the kind of the participial modification clause can be accurately decided.



5: part of speech matrix table, a: Japanese language text, b: input, c: output, d: analysis result

(54) INFORMATION RETRIEVER

(11) 4-184678 (A) (43) 1.7.1992 (19) JP

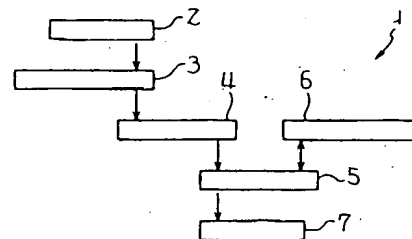
(21) Appl. No. 2-316504 (22) 20.11.1990

(71) RICOH CO LTD (72) SHIYOU IMASATO

(51) Int. Cl. G06F15/40

PURPOSE: To dispense with a key work, etc., and to facilitate retrieval work by performing clustering on hierarchical structure based on similarity between knowledge information, and retrieving the hierarchical structure from a high-order one to a low-order one according to an operation input.

CONSTITUTION: A similarity detecting means 2 calculates the similarity of arbitrary two pieces of knowledge information by prescribed processing. Thence, a clustering means 3 assumes each knowledge information as one cluster as an initial state, and detects and unifies a pair of clusters with the highest similarity, and calculates the similarity again simultaneously with such connection. Tree-structured hierarchical structure is formed, and is stored in a cluster storage means 4. Thence, a cluster selection mean 5 traces the cluster in the means 4 from the high-order one to the low-order one according to the operation input, and retrieves the knowledge information desired by a worker. At this time, an auxiliary information offering means 6 offers auxiliary information to clarify an operation. By employing such constitution, it is possible to dispense with the character-input of a retrieval condition by the operation of a keyboard, and to facilitate the retrieval work, and to improve working efficiency.



7: information output means

(54) OPTICAL DISK STORAGE IMAGE DOCUMENT RETRIEVER

(11) 4-184679 (A) (43) 1.7.1992 (19) JP

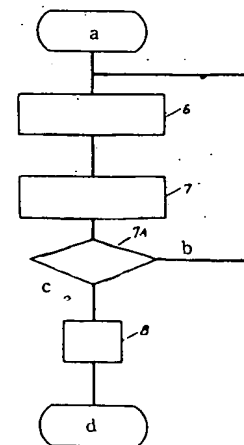
(21) Appl. No. 2-317288 (22) 20.11.1990

(71) MATSUSHITA ELECTRIC IND CO LTD (72) KOJI YANO

(51) Int. Cl. G06F15/40, G11B27/00, H04N5/85

PURPOSE: To perform the copy of an image document with password with high efficiency as sufficiently keeping security by enabling a retrieval instruction to be issued in spite of the presence/absence of a password when the image document targeted to copy is retrieved.

CONSTITUTION: The image document targeted to copy selected at a selection processing part 6 is retrieved by a retrieval processing part 7. In such a case, processing by software to automatically attach the same password as that of the image document targeted to copy on the image document stored in a device main body is performed at the processing part 7 by supplying an instruction capable of performing retrieval in spite of the presence/absence of the password, and the image document with password is retrieved. Thence, when it is judged that a retrieval result is normal by a judging means 7A, a retrieved image document is registered on a registration processing part 8 newly. Thereby, it is possible to copy the image document with password with high efficiency without inputting the password by a user.



a: copy start, b: abnormal completion(NO), c: normal completion(YES), d: copy end

⑩ 日本国特許庁(JP)

⑪ 特許出願公開

⑫ 公開特許公報(A) 平4-184678

⑬ Int.Cl.⁵

識別記号

庁内整理番号

⑭ 公開 平成4年(1992)7月1日

G 06 F 15/40

5 1 0 J

7056-5L

審査請求 未請求 請求項の数 8 (全11頁)

⑮ 発明の名称 情報検索装置

⑯ 特 願 平2-316504

⑰ 出 願 平2(1990)11月20日

⑱ 発 明 者 今 郷 昭 東京都大田区中馬込1丁目3番6号 株式会社リコー内
⑲ 出 願 人 株 式 会 社 リ コ ー 東京都大田区中馬込1丁目3番6号
⑳ 代 理 人 弁 理 士 柏 木 明

明 細 書

1. 発明の名称 情報検索装置

2. 特許請求の範囲

1. 予め複数の知識情報が記憶された情報記憶手段と、前記情報記憶手段内の知識情報間の類似度を算定する類似度検出手段と、この類似度検出手段の算定結果に従って前記知識情報を複数のクラスタからなる階層構造にクラスタリングするクラスタリング手段と、この情報クラスタリング手段がクラスタリングした前記知識情報の階層構造を記憶するクラスタ記憶手段と、このクラスタ記憶手段の記憶内容や前記知識情報の内容に従って少なくとも前記クラスタの階層構造を内包する補助情報を提示する補助情報提供手段と、操作入力に従って前記クラスタを階層構造の上位から下位に向かって順次検索するクラスタ選択手段と、このクラスタ選択手段が検索したクラスタ下の前記

知識情報を出力する情報出力手段とよりなることを特徴とする情報検索装置。

2. 予め複数の知識情報が登録キーワードと共に記憶された情報記憶手段を設け、前記登録キーワードの一致率に従って前記知識情報の類似度を算定する類似度検出手段を設けたことを特徴とする請求項1記載の情報検索装置。

3. 各々少なくとも一つの単語を内包する複数の知識情報が予め記憶された情報記憶手段を設け、前記単語の一致率に従って前記知識情報の類似度を算定する類似度検出手段を設けたことを特徴とする請求項1記載の情報検索装置。

4. 各々少なくとも一つの文字を内包する複数の知識情報が予め記憶された情報記憶手段を設け、前記文字の一致率に従って前記知識情報の類似度を算定する類似度検出手段を設けたことを特徴とする請求項1記載の情報検索装置。

5. 各知識情報が一つのクラスタに付属するよ

うに規制するクラスタ限定手段を設けたことを特徴とする請求項 1, 2, 3 又は 4 記載の情報検索装置。

6. クラスタ選択手段が検索する知識情報の個数を指定する検索個数指定手段を設け、この検索個数指定手段の指定個数の近似値に各クラスタに付属する前記知識情報の個数を規制するクラスタ限定手段を設けたことを特徴とする請求項 5 記載の情報検索装置。

7. クラスタ選択手段が検索する知識情報の個数を指定する検索個数指定手段を設け、この検索個数指定手段の指定個数以下に各クラスタに付属する前記知識情報の個数を規制するクラスタ限定手段を設けたことを特徴とする請求項 5 記載の情報検索装置。

8. 予め情報記憶手段に記憶された複数の知識情報の一つを検索基準として指定する情報指定手段を設け、この情報指定手段で指定された知識情

報が付属するクラスタの一つを選択するクラスタ選択手段を設けたことを特徴とする請求項 1, 2, 3 及び 4 記載の情報検索装置。

3. 発明の詳細な説明

産業上の利用分野

本発明は、データベース検索システムなどに利用される情報検索装置に関するものである。

従来の技術

一般的な情報検索装置は、予め複数の知識情報に対応する登録キーワードと共に情報記憶手段に格納しておき、入力する検索キーワードと登録キーワードとのマッチングで知識情報を検索するようになっている。そして、このような情報検索装置は、現在では検索精度を向上させるために検索キーワードと演算子とを組合わせて検索条件を生成するようになっており、その具体例としては、「化学・薬学・生物医学のオンライン情報検索」

(佐本光雄、坂本徹朗、黒沢正彦共著 地人書館 1981年)に開示された情報検索方法がある。

その一例である組立型の検索では、最初に単独の検索キーワードで検索作業を行ってから、結果を演算子で組合わせるようになっており、例えば、検索キーワードを A～C とすると、

1, select A

(A に対応する情報検索)

2, select B

(B に対応する情報検索)

3, combine 1 and 2

(1 と 2 の検索結果の共通集合を検出)

と云うように検索条件を複数回で入力する。

また、他の例である試行錯誤型の検索では、最初に検索条件式を組立てて検索結果を確認しながら検索条件式を修正するようになっており、例えば、検索キーワードを A～C とすると、

A or B and C

と云うような検索条件を一回で入力する。

発明が解決しようとする課題

上述のような情報検索方法は、情報検索の専任者の使用を想定しており、初心者には検索条件の選定や生成が容易でない。

具体的には、知識情報の検索に適した検索キーワードを作業者が考える必要があるので、個人差により結果が異なることになる。このような課題は、利用できる検索キーワードが予め設定されている統制キーワード方式の情報検索装置では発生しないが、これは作業の自由度が低い。そして、作業者の所望により検索キーワードを入力する自由キーワード方式の情報検索装置では、上述のような課題が発生すると共に、初心者が適当に検索キーワードを生成しても良好な検索結果が得られないことが多発する。

また、上述のような情報検索装置では、検索キーワードや演算子等からなる検索条件をキーボー

ドによる文字入力で行なうようになっており、その操作が煩雑で作業能率を低下させている。

さらに、上述した情報検索方法では、複数の演算子を内包する検索条件式を入力できるようになっているが、この場合は使用者が演算子の優先順位を認識している必要がある。つまり、

A or B and C

と云う検索条件式を入力した場合、これが

(A or B) and C

A or (B and C)

の何れになるかは情報検索装置の設定によるので、使用者の要望に反した検索結果が検出されることがある。

例えば、初心者が情報検索装置を使用する場合の作業を想定すると、まず、思いついた一つの検索キーワードを入力し、検索結果が多い場合は他の検索キーワードとのandをとり、検索結果が少ない場合は他の検索キーワードとのorをとる。そ

文書として入力し、文書情報を検索すると云うものである。

つまり、データベース内の文書情報の各々にキー文書に対するキーワード包含率を、(キー文書と一致したキーワード数)/(キー文書のキーワード数)として算出し、このキーワード包含率が高い文書情報について係受け関係の類似性とキーワードの重みとでキー文書との類似度を算出し、この類似度の高いものから文書情報を作業者に提示するようになっている。

しかし、これではキー文書と文書情報とを互いの一部で比較しているので、良好な検索結果を得ることが困難である。さらに、このようにして検索した文書情報を類似度に従って順番に提示するのでは、作業者の要望に最も近い文書情報が下位に分散される懸念があり、最終的には作業者が文書情報の内容を確認する必要があるなどして作業能率が低下することになる。

して、もしも検索結果が予想に反するものである場合は、検索結果を放棄して作業を最初から繰返すことになる。しかし、これでは操作が煩雑で作業能率が低く、良好な検索結果を得ることが困難である。

上述のような情報検索装置を使用する際は、一般的に作業者は適正な個数の検索結果を要望している場合が多く、検索意図に適した検索結果が得られても個数が過多であったり過少であった場合は、検索意図を狭くしたり拡大するなどして操作を繰返している。しかし、これでは操作が煩雑で作業能率が低く、操作を繰返すと共に検索条件の生成が困難となる。

ここで、上述のような検索キーワードの入力を要しない情報検索手段の一例が、稲垣博人“文書内容検索システムにおける現状と課題”情報管理 vol.33 No.2 1980 に開示されている。これは、検索キーワードの代わりに所定の文書情報をキー

課題を解決するための手段

請求項1記載の発明は、予め複数の知識情報が記憶された情報記憶手段を設け、情報記憶手段内の知識情報間の類似度を算定する類似度検出手段を設け、この類似度検出手段の算定結果に従って知識情報を複数のクラスタからなる階層構造にクラスタリングするクラスタリング手段を設け、この情報クラスタリング手段がクラスタリングした知識情報の階層構造を記憶するクラスタ記憶手段を設け、このクラスタ記憶手段の記憶内容や知識情報の内容に従って少なくともクラスタの階層構造を内包する補助情報を提示する補助情報提供手段を設け、操作入力に従ってクラスタを階層構造の上位から下位に向かって順次検索するクラスタ選択手段を設け、このクラスタ選択手段が検索したクラスタ下の知識情報を出力する情報出力手段を設けた。

請求項2記載の発明は、予め複数の知識情報が

登録キーワードと共に記憶された情報記憶手段を設け、登録キーワードの一致率に従って知識情報の類似度を算定する類似度検出手段を設けた。

請求項3記載の発明は、各々少なくとも一つの単語を内包する複数の知識情報が予め記憶された情報記憶手段を設け、単語の一致率に従って知識情報の類似度を算定する類似度検出手段を設けた。

請求項4記載の発明は、各々少なくとも一つの文字を内包する複数の知識情報が予め記憶された情報記憶手段を設け、文字の一致率に従って知識情報の類似度を算定する類似度検出手段を設けた。

請求項5記載の発明は、各知識情報が一つのクラスタに付属するように規制するクラスタ限定手段を設けた。

請求項6記載の発明は、クラスタ選択手段が検索する知識情報の個数を指定する検索個数指定手段を設け、この検索個数指定手段の指定個数の近似値に各クラスタに付属する知識情報の個数を規

制するクラスタ限定手段を設けた。

請求項7記載の発明は、クラスタ選択手段が検索する知識情報の個数を指定する検索個数指定手段を設け、この検索個数指定手段の指定個数以下に各クラスタに付属する知識情報の個数を規制するクラスタ限定手段を設けた。

請求項8記載の発明は、予め情報記憶手段に記憶された複数の知識情報の一つを検索基準として指定する情報指定手段を設け、この情報指定手段で指定された知識情報が付属するクラスタの一つを選択するクラスタ選択手段を設けた。

作用

請求項1記載の発明は、情報記憶手段が予め複数の知識情報を記憶しており、これらの知識情報間の類似度を類似度検出手段が算定し、この算定結果に従って知識情報を複数のクラスタからなる階層構造にクラスタリング手段がクラスタリングし、このクラスタリングした知識情報の階層構造

をクラスタ記憶手段が記憶し、この記憶内容や知識情報の内容に従って少なくともクラスタの階層構造を内包する補助情報を補助情報提供手段が提示し、操作入力に従ってクラスタを階層構造の上位から下位に向かってクラスタ選択手段が順次検索し、この検索されたクラスタ下の知識情報を情報出力手段が出力することで、作業者は検索作業を行なう場合に二つのクラスタの一方を選択するだけなので、作業者がキーワードや演算子等からなる検索条件を考えたり、このような検索条件をキーボードの操作で文字入力する必要や、試行錯誤に従って検索作業を繰返す必要がなく、さらに、予め知識情報に登録キーワードを設定しておくことも要しない。

請求項2記載の発明は、予め複数の知識情報と共に記憶された登録キーワードの一致率に従って類似度検出手段が知識情報の類似度を算定することで、簡易に知識情報の類似度を算定することが

できる。

請求項3記載の発明は、各々少なくとも一つの単語を内包する複数の知識情報の単語の一致率に従って類似度検出手段が知識情報の類似度を算定することで、簡易に知識情報の類似度を算定することができ、しかも、予め知識情報に登録キーワードを設定しておくことを要しない。

請求項4記載の発明は、各々少なくとも一つの文字を内包する複数の知識情報の文字の一致率に従って類似度検出手段が知識情報の類似度を算定することで、簡易に知識情報の類似度を算定することができ、しかも、予め知識情報に登録キーワードを設定しておくことを要しない。

請求項5記載の発明は、各知識情報が一つのクラスタに付属するようにクラスタ限定手段が規制することで、クラスタ選択手段の動作を制御するなどして検索結果の個数を作業者の所望の個数に近似させることができる。

請求項 6 記載の発明は、クラスタ選択手段が検索する知識情報の個数を検索個数指定手段で指定し、この指定個数の近似値に各クラスタに付属する知識情報の個数をクラスタ限定手段が規制することで、作業者は略所望する個数の検索結果を得ることができる。

請求項 7 記載の発明は、クラスタ選択手段が検索する知識情報の個数を検索個数指定手段で指定し、この指定個数以下に各クラスタに付属する知識情報の個数をクラスタ限定手段が規制することで、作業者は略所望する個数の検索結果を得ることができる。

請求項 8 記載の発明は、予め情報記憶手段に記憶された複数の知識情報の一つを検索基準として情報指定手段で指定し、この指定された知識情報が付属するクラスタの一つをクラスタ選択手段で選択することで、作業者がキーワードや演算子等からなる検索条件を考える必要がなく、また、類

似度が高い知識情報を瞬接して提示することができる。

実施例

請求項 1～4 記載の発明の実施例を第 1 図ないし第 5 図に基づいて説明する。まず、本実施例の情報検索装置 1 は、第 1 図に例示するように、予め複数の知識情報が記憶された情報記憶手段(図示せず)に、前記知識情報間の類似度を算定する類似度検出手段 2 が接続されており、この類似度検出手段 2 に前記知識情報をクラスタリングするクラスタリング手段 3 が接続されている。そして、このクラスタリング手段 3 には前記知識情報の階層構造を記憶するクラスタ記憶手段 4 が接続されており、このクラスタ記憶手段 4 には操作入力に従って前記クラスタを階層構造の上位から下位に向かって順次検索するクラスタ選択手段 5 が接続されている。そして、このクラスタ選択手段 5 には、前記知識情報に対する前記クラスタの階層構

造を提示する補助情報提供手段 6 と、前記知識情報を表示や印刷で出力する情報出力手段 7 とが接続されている。

ここで、本発明の情報検索装置 1 で利用するクラスタリングとは、所定の情報を複数のクラスタからなる階層構造に分類すると云う既存の概念であり、例えば、河合至簡“多変量解析入門Ⅱ”森北出版 数学ライブラリー 46、1978 などに詳述されている。

このような構成において、本実施例の情報検索装置 1 の各部の動作を以下に詳述する。

まず、類似度検出手段 2 は、任意の二つの知識情報の類似度を所定の演算処理で算定するようになっており、この演算手段としては以下に例示するようなものがある。

第一に例示する手段は、請求項 2 記載の発明に記載したもので、予め知識情報が登録キーワードと共に記憶されている 合に、登録キーワードの

一致率に従って知識情報の類似度を算定する。つまり、二つの知識情報 A、B の登録キーワードの個数が N_A 、 N_B であり、これらに共通する登録キーワードの個数を N_{AB} とすると、

$$\text{類似度} = N_{AB} / (N_A + N_B - N_{AB})$$

とする。このようにすることで、この類似度は二つの知識情報 A、B に共通する登録キーワードがなければ 0 となり、全ての登録キーワードが共通ならば 1 となる。

つぎに、第二に例示する手段は、請求項 3 記載の発明に記載したもので、文書等からなる知識情報が各々少なくとも一つの単語を内包する場合、これらの単語の一致率に従って知識情報の類似度を算定する。つまり、二つの知識情報 A、B に出現する単語の個数が N_A 、 N_B で共通する単語の個数を N_{AB} とすると、

$$\text{類似度} = N_{AB} / (N_A + N_B - N_{AB})$$

とする。このようにすることで、この類似度は二

つの知識情報 A, B に共通する単語がなければ 0 となり、全ての単語が共通ならば 1 となる。なお、この方法を実施する場合は、助詞や助動詞等の付属語を除去してから類似度を算定することや、名詞だけを抽出して類似度を算定することなどが考えられる。

さらに、第三に例示する手段は、請求項 4 記載の発明に記載したもので、文書や単語等からなる知識情報が各々少なくとも一つの文字を内包する場合、これらの文字の一致率に従って知識情報の類似度を算定する。この場合、最初に所定の文字 C の知識情報内における使用率を、

$$\text{使用率} = (\text{文字 C の個数}) / (\text{文字の総数})$$

として算定し、それぞれの文字の使用率を要素とするベクトルの内積で二つの知識情報の類似度を算定する。なお、この方法を実施する場合は、主に付属語として利用される平仮名を除去してから類似度を算定することなどが考えられる。

なお、この再計算の方法としては、併合されるクラスタを Cold₁, Cold₂ として併合後のクラスタを Cnew とし、これに併合するクラスタを Celse とすると、Cnew と Celse との類似度を、Cold₁ と Celse、Cold₂ と Celse の類似度のうち大きい方、平均値、小さい方等とすることなどが考えられる。

そして、併合するクラスタがなくなるまで上述の作業を繰返すことで、クラスタリング手段 3 は複数の知識情報から階層構造を形成する。

また、クラスタ記憶手段 4 は、上述のようにしてクラスタリング手段 3 でクラスタリングされた知識情報の木構造状の階層構造を記憶する。なお、第 2 図に例示した階層構造では、最初に知識情報 ②, ③ が併合されて次に知識情報 ⑦, ⑧ が併合されたことが分かる。実際には、クラスタリング記憶手段 4 内の記憶構造は、第 3 図に例示するような形態となっており、所定のクラスタを指定する

また、第四に例示する手段は、知識情報が画像のイメージデータ等からなる場合に有効なもので、抽出した特徴ベクトルの内積で二つの知識情報の類似度を算定する。

上述のように知識情報の種別に従って各種の演算処理を行なうことで、類似度検出手段 2 は二つの知識情報の類似度を 0~1 の範囲で算定することができる。

つぎに、クラスタリング手段 3 は、初期状態として各知識情報を要素が一つのクラスタと見なし、第 2 図に例示するように、類似度が高いクラスタ対から順次併合して木構造状の階層構造を形成する。そこで、その動作を簡易に詳述する。

まず、全ての知識情報を各々クラスタとする。なお、この段階では各クラスタの類似度は知識情報の類似度と同一である。

つぎに、類似度が最も高いクラスタ対を検出して併合し、この連結と共に類似度を再計算する。

と全ての子クラスタを検出することができる。

さらに、補助情報提供手段 6 は、初心者でも操作が簡明であるように検索作業の補助情報を提供するもので、実際には上述のようなクラスタ記憶手段 4 の記憶内容や知識情報の内容に従って階層構造に基づいたクラスタの関係を内包した補助情報を表示や印刷で提示する。さらに、このような補助情報としては、以下に例示するようなことを付加することが実施可能である。

1. 知識情報に表題がある場合、各クラスタ下の知識情報の表題の一覧を提示する。
2. 知識情報に表題がある場合、表題を構成する単語を出現数の多い順に提示する。
3. 各知識情報の内容の一部を一覧提示する。
4. 知識情報にキーワードがある場合は、各キーワードを対応情報が多い順に提示する。
5. 知識情報をランダムに選出して内容を提示する。

このようにすることで、作業者は補助情報提供手段 6 が提示した内容に基づいて検索作業を容易に行なうことができる。

そして、クラスタ選択手段 5 は、操作入力に従ってクラスタ記憶手段 4 内のクラスタを上位から下位に辿り、作業者が所望する知識情報を検索するようになっていく。つまり、注目点を最上位のクラスタとして二つの子クラスタに関する補助情報を補助情報提供手段 6 で提示し、作業者の選択による操作入力に従って次の二つの子クラスタの補助情報を提示する。そして、このような動作を繰返すことで、作業者は所望の知識情報を検出することができる。ここで、上述のような補助情報の一部としてクラスタに付属する知識情報の個数を提示することで、作業者は所望の個数の知識情報を得ることができる。

さらに、情報出力手段 7 は、クラスタ選択手段 5 で検索された知識情報を表示や印刷等で出力す

るようになっており、ここでは知識情報をクラスタリングされた順に出力するようになっている。例えば、第 2 図及び第 3 図に例示した階層構造では、クラスタ ⑤ は知識情報 ④ とクラスタ ③ とからなり、このクラスタ ③ は知識情報 ⑤、⑥ からなるので、クラスタ ⑤ には知識情報 ⑤、⑥ が取込まれてから知識情報 ④ が取込まれたことになる。そこで、この場合の情報出力手段 7 の出力順序は、知識情報 ⑤ → 知識情報 ⑥ → 知識情報 ④ となるか、

知識情報 ⑥ → 知識情報 ⑤ → 知識情報 ④ となる。

そこで、この情報検索装置 1 で知識情報をクラスタリングする場合の処理動作を、第 4 図に例示するフローチャートに基づいて以下に説明する。まず、予め情報記憶手段内に記憶された全ての知識情報に対して類似度を類似度検出手段 2 が算定する。そこで、クラスタリング手段 3 は、各知識

情報をクラスタと見なして最も類似度が高いクラスタ対を選出し、このクラスタ対を新たなクラスタとして類似度を再計算し、併合するクラスタが無くなるまで類似度に従ったクラスタの併合を繰返す。

このようにすることで、知識情報が階層構造にクラスタリングされる。

つぎに、この情報検索装置 1 でクラスタリングされた知識情報を検索する場合の処理動作を、第 5 図に例示するフローチャートに基づいて以下に説明する。まず、クラスタリング記憶手段 4 の記憶内容に従って機器が最上位のクラスタに注目するので、作業者は二つの子クラスタの補助情報を補助情報提供手段 6 から得て一方を選択する。そこで、選択されたクラスタの二つの子クラスタの補助情報が得られるので、上述の操作を繰返して所望する知識情報を検索し、この検索された知識情報の内容を情報出力手段 7 で出力することで検

索作業が完了する。

上述のように、この情報検索装置 1 では、検索作業を行なう場合に二つのクラスタの一方を選択するだけなので、作業者がキーワードや演算子等からなる検索条件を考えたり、このような検索条件をキーボードの操作で文字入力する必要がなく、試行錯誤に従って検索作業を繰返す必要もないので、検索作業の操作が容易で作業能率が良好である。また、この情報検索装置 1 では、既存の技術で知識情報がクラスタリングされ、予め知識情報に登録キーワードを設定しておくことを要しないので、機器の構造が簡易で形成が容易である。

また、請求項 2 記載の発明に記載したように、知識情報が登録キーワードと共に記憶されている場合は、登録キーワードの一致率に従って知識情報の類似度を算定することができ、請求項 3 記載の発明に記載したように、文書等からなる知識情報が単語を内包する場合は、これらの単語の一致

率に従って知識情報の類似度を算定することができる。請求項 4 記載の発明に記載したように、文や単語等からなる知識情報が文字を内包する場合は、これらの文字の一致率に従って知識情報の類似度を算定することができる。このようにすることで、容易に知識情報の類似度を算定して情報検索装置 1 を実現することができる。

さらに、請求項 5～7 記載の発明の実施例を図 6 図及び第 7 図に基づいて説明する。まず、この情報検索装置 8 では、第 6 図に例示するように、クラスタ記憶手段 4 とクラスタ選択手段 5 との間に、各知識情報が一つのクラスタに付属するように規制するクラスタ限定手段 9 が設けられており、このクラスタ限定手段 9 に検索する知識情報の個数を指定する検索個数指定手段 10 が接続されている。なお、この他の構造は前述の情報検索装置 1 と同様になっている。

このような構成において、本実施例の情報検索

装置 8 の各部の動作を以下に詳述する。なお、前述の情報検索装置 1 と同一の部分 2～7 は同様に機能する。

まず、検索個数指定手段 10 は、作業者が要望する検索結果の個数が入力されるようになっており、この入力値に従ってクラスタ限定手段 9 の動作が制御されるようになっている。

そして、このクラスタ限定手段 9 は、各知識情報が一つのクラスタに付属するように規制すると共に、クラスタ選択手段 5 の動作を制御するなどして検索結果の個数が検索個数指定手段 10 で指定された個数以下の最大数や近似値となるようにクラスタを限定する。

つまり、第 7 図のフローチャートに例示するように、まず、検索個数指定手段 10 で検索を要望する知識情報の個数を指定すると、注目点を最上位のクラスタとして付属する知識情報の個数である検索個数が指定個数と比較される。そこで、こ

の指定個数より検索個数が小さければ二つのクラスタの一方が選択され、この選択されたクラスタにも同様な処理が行なわれて検索個数指定手段 10 の指定個数以下の最大数や近似した個数の知識情報が検出される。

このようにすることで、作業者は略所望する個数の検索結果を得ることができるので、検索結果が所望の個数になるまで試行錯誤に従って検索作業を繰返すような必要がない。そして、上述のように検索個数指定手段 10 で大きい個数を指定することで、漏れが少ない検索結果を得ることができ、小さい個数を指定することで、不要な情報量が少ない検索結果を得ることができる。

つぎに、請求項 8 記載の発明の実施例を図 8 図及び第 9 図に基づいて説明する。まず、この情報検索装置 11 では、第 8 図に例示するように、検索基準となる知識情報を指定する情報指定手段 12 がクラスタ選択手段 5 に接続されている。なお、

この他の構造は前述の情報検索装置 1 と同様になっている。

このような構成において、本実施例の情報検索装置 11 の情報指定手段 12 に、検索基準となる文書情報等の知識情報を入力すると、この知識情報が付属するクラスタの一つをクラスタ選択手段 5 が選択するようになっている。なお、上述のような検索基準となる知識情報は、予め情報記憶手段内に格納されている知識情報の一つであり、これを得るためには表題や検索キーワードによる検索作業を行なうことが考えられる。また、新たに知識情報を登録した場合は、これを検索基準とすることも可能である。

そこで、この情報検索装置 11 で検索作業を行なう場合の動作を図 9 図に例示するフローチャートに基づいて以下に説明する。まず、情報指定手段 12 で指定された知識情報が付属する最下位のクラスタ対の一方がクラスタ選択手段 5 で検出さ

れ、このクラスタの補助情報が補助情報提供手段6から作業者に出力される。そこで、この作業者は、そのクラスタに付属する知識情報を情報出力手段7から出力させて検索作業を完了するか、注目点を上位のクラスタと対のクラスタとの一方に移動させるかを判断する。そこで、注目点を移動する場合は、そのクラスタに対して上述と同様な作業を行なうことで、作業者が所望する検索結果が得られることになる。

このようにすることで、この情報検索装置11では、検索基準として知識情報を指定すれば良いので、作業者がキーワードや演算子等からなる検索条件を考える必要がなく、また、類似度が高い知識情報を簡易して提示することができるので、作業者の要望に近い知識情報を簡易に得ることができ、検索作業の操作が容易で作業能率が良好である。

発明の効果

条件を考えたり、このような検索条件をキーボードの操作で文字入力する必要がなく、試行錯誤に従って検索作業を繰返す必要もないので、検索作業の操作が容易で作業能率が良好であり、予め知識情報に登録キーワードを設定しておくことを要しないので、機器の構造が簡易で形成が容易である等の効果を有するものである。

請求項2記載の発明は、予め複数の知識情報が登録キーワードと共に記憶された情報記憶手段を設け、登録キーワードの一致率に従って知識情報の類似度を算定する類似度検出手段を設けたことにより、簡易に知識情報の類似度を算定することができ、高性能な情報検索装置を容易に実施することができる等の効果を有するものである。

請求項3記載の発明は、各々少なくとも一つの単語を内包する複数の知識情報が予め記憶された情報記憶手段を設け、単語の一致率に従って知識情報の類似度を算定する類似度検出手段を設けた

請求項1記載の発明は、予め複数の知識情報が記された情報記憶手段を設け、情報記憶手段内の知識情報間の類似度を算定する類似度検出手段を設け、この類似度検出手段の算定結果に従って知識情報を複数のクラスタからなる階層構造にクラスタリングするクラスタリング手段を設け、この情報クラスタリング手段がクラスタリングした知識情報の階層構造を記憶するクラスタ記憶手段を設け、このクラスタ記憶手段の記憶内容や知識情報の内容に従って少なくともクラスタの階層構造を内包する補助情報を提示する補助情報提供手段を設け、操作入力に従ってクラスタを階層構造の上位から下位に向かって順次検索するクラスタ選択手段を設け、このクラスタ選択手段が検索したクラスタ下の知識情報を出力する情報出力手段を設けたことにより、作業者は検索作業を行なう場合に二つのクラスタの一方を選択するだけで、作業者がキーワードや演算子等からなる検索

ことにより、簡易に知識情報の類似度を算定することができ、しかも、予め知識情報に登録キーワードを設定しておくことを要しないので、高性能な情報検索装置を容易に実施することができる等の効果を有するものである。

請求項4記載の発明は、各々少なくとも一つの文字を内包する複数の知識情報が予め記憶された情報記憶手段を設け、文字の一致率に従って知識情報の類似度を算定する類似度検出手段を設けたことにより、簡易に知識情報の類似度を算定することができ、しかも、予め知識情報に登録キーワードを設定しておくことを要しないので、高性能な情報検索装置を容易に実施することができる等の効果を有するものである。

請求項5記載の発明は、各知識情報が一つのクラスタに付属するように規制するクラスタ限定手段を設けたことにより、クラスタ選択手段の動作を制御するなどして検索結果の個数を作業者の所

望の個数に近似させることができ、高性能な情報検索装置を容易に実施することができる等の効果を有するものである。

請求項 6 記載の発明は、クラスタ選択手段が検索する知識情報の個数を指定する検索個数指定手段を設け、この検索個数指定手段の指定個数の近似値に各クラスタに付属する知識情報の個数を規制するクラスタ限定手段を設けたことにより、作業者は略所望する個数の検索結果を得ることができるので、検索結果が所望の個数になるまで試行錯誤に従って検索作業を繰返すような必要がなく、大きい個数を指定して漏れが少ない検索結果を得たり、小さい個数を指定して不要な情報量が少ない検索結果を得るようなことができる等の効果を有するものである。

請求項 7 記載の発明は、クラスタ選択手段が検索する知識情報の個数を指定する検索個数指定手段を設け、この検索個数指定手段の指定個数以下

に各クラスタに付属する知識情報の個数を規制するクラスタ限定手段を設けたことにより、作業者は略所望する個数の検索結果を得ることができるので、検索結果が所望の個数になるまで試行錯誤に従って検索作業を繰返すような必要がなく、大きい個数を指定して漏れが少ない検索結果を得たり、小さい個数を指定して不要な情報量が少ない検索結果を得るようなことができる等の効果を有するものである。

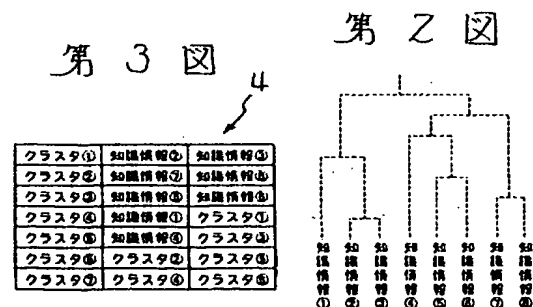
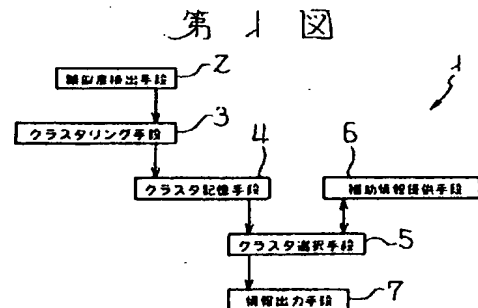
請求項 8 記載の発明は、予め情報記憶手段に記憶された複数の知識情報の一つを検索基準として指定する情報指定手段を設け、この情報指定手段で指定された知識情報が付属するクラスタの一つを選択するクラスタ選択手段を設けたことにより、検索基準として知識情報を指定すれば良いので、作業者がキーワードや演算子等からなる検索条件を考える必要がなく、また、類似度が高い知識情報を關接して提示することができるので、作業者

の要望に近い知識情報を簡易に得ることができ、検索作業の操作が容易で作業能率が良好である等の効果を有するものである。

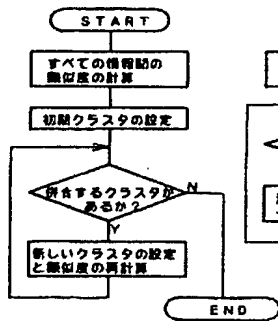
4. 図面の簡単な説明

図面は本発明の実施例を示すものであり、第 1 図は請求項 1～4 記載の発明のブロック図、第 2 図及び第 3 図はデータ構造の概念説明図、第 4 図及び第 5 図はフローチャート、第 6 図は請求項 5～7 記載の発明のブロック図、第 7 図はフローチャート、第 8 図は請求項 8 記載の発明のブロック図、第 9 図はフローチャートである。

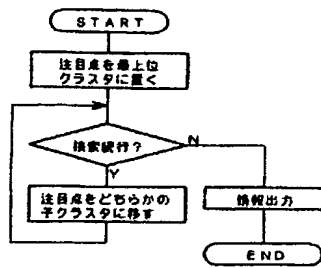
1、8、11…情報検索装置、2…類似度検出手段、3…クラスタリング手段、4…クラスタ記憶手段、5…クラスタ選択手段、6…補助情報提供手段、7…情報出力手段、9…クラスタ限定手段、10…検索個数指定手段、12…情報指定手段



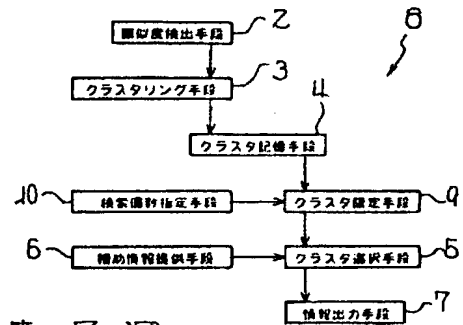
第 4 図



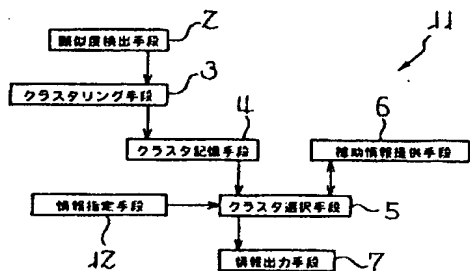
第 5 図



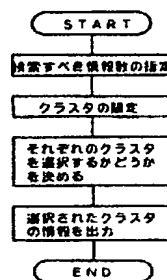
第 6 図



第 8 図



第 7 図



第 9 図

